OBJN
Online Brazilian Journal of Nursing
ISSN: 1676 - 4285 | www.objnursing.uff.br

Envío: 28-May-2026
Aprobado: 26-Ago-2026

SELECCIÓN DE ESTUDIOS ASISTIDA POR LA INTELIGENCIA ARTIFICIAL CLAUDE AI: UNA GUÍA METODOLÓGICA

SELEÇÃO DE ESTUDOS ASSISTIDA PELA INTELIGÊNCIA ARTIFICIAL *CLAUDE AI*: UM GUIA METODOLÓGICO

STUDY SELECTION ASSISTED BY CLAUDE AI ARTIFICIAL INTELLIGENCE: A METHODOLOGICAL GUIDE

Rômulo Severo Sampaio[1] ORCID: 0000-0002-0657-2646
Yasmim Ribeiro Fracaroli[2] ORCID: 0009-0004-2400-2323
Anna Paula Mendes Marques de Lima Franco[2] ORCID: 0009-0009-8847-7914
Marianne Alves Rodrigues[2] ORCID: 0009-0001-6504-9089
Francini Castilha do Nascimento[2] ORCID: 0000-0001-7860-0493
Cecília Santos de Lima[2] ORCID: 0009-0004-3706-2368
Isabelle Cristinne Pinto Sampaio Costa[2] ORCID: 0000-0002-2611-8643

[1] Universidade Professor Edson Antônio Velano. Alfenas, Minas Gerais, Brasil.
[2] Universidade Federal de Alfenas (UNIFAL-MG). Alfenas, Minas Gerais, Brasil.

Cómo citar: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, et al. Study selection assisted by Claude AI artificial intelligence: a methodological guide. Online Braz J Nurs. 2026;25(Suppl1):e20267039. https://doi.org/10.17665/1676-4285.20267039

RESUMO

Objetivo: Apresentar um guia metodológico para a seleção de estudos em revisões de literatura com o auxílio do *Claude AI*. **Método:** Estudo de reflexão teórico-metodológica, fundamentado em revisão narrativa da literatura e na experiência empírica dos autores na condução de sínteses de evidências. **Resultados:** Propõe-se um fluxo metodológico estruturado em cinco etapas: (1) exportação e deduplicação inicial no *EndNote*; (2) deduplicação adicional e fragmentação do corpus em blocos anuais no *Rayyan*; (3) criação de um projeto no *Claude AI*, incluindo os critérios de elegibilidade; (4) triagem de títulos e resumos assistida por inteligência artificial (IA), com geração automatizada de planilha contendo a decisão e a justificativa para cada registro; e (5) retorno ao *Rayyan* para revisão humana independente e posterior integração ao fluxograma PRISMA. Esse processo configura um modelo teórico-operacional que fornece recomendações práticas para a definição dos critérios de elegibilidade, elaboração estruturada de prompts e registro transparente das decisões. **Conclusão:** O guia estabelece uma estrutura inovadora para a incorporação da IA na etapa de seleção de estudos em revisões de literatura. Essa abordagem acelera as fases iniciais da síntese de evidências e orienta o uso responsável da tecnologia. Como limitação, aponta-se a ausência de validação empírica multicêntrica do fluxo proposto, o que demanda sua aplicação e avaliação em pesquisas futuras.

Descritores: Inteligência Artificial; Inteligência Artificial Generativa; Prática Clínica Baseada em Evidências; Metodologia como Assunto; Pesquisa; Revisão Sistemática.

ABSTRACT

Objective: To present a methodological guide for study selection in literature reviews assisted by Claude AI. **Method:** A theoretical-methodological reflection study grounded in a narrative literature review and the authors' empirical experience in conducting evidence syntheses. **Results:** A five-stage structured methodological workflow is proposed: (1) export and initial deduplication in EndNote; (2) additional deduplication and corpus segmentation into annual blocks in Rayyan; (3) creation of a project in Claude AI incorporating the eligibility criteria; (4) AI-assisted title and abstract screening with automated generation of a spreadsheet containing the decision and justification for each record; and (5) return to Rayyan for independent human review and subsequent integration into the PRISMA flow diagram. This process establishes a theoretical-operational model that provides practical recommendations for defining eligibility criteria, developing structured prompts, and transparently documenting decisions. **Conclusion:** The guide establishes an innovative framework for incorporating AI into the study selection phase of literature reviews. This approach accelerates the initial stages of evidence synthesis and promotes the responsible use of technology. A noted limitation is the lack of multicenter empirical validation of the proposed workflow, which warrants its application and evaluation in future research.

Descriptors: Artificial Intelligence; Generative Artificial Intelligence; Evidence-Based Practice; Methodology as a Subject; Research; Systematic Review.

RESUMEN

Objetivo: Presentar una guía metodológica para la selección de estudios en revisiones de literatura con el auxilio de Claude AI. **Método:** Estudio de reflexión teórico-metodológica, fundamentado en una revisión narrativa de la literatura y en la experiencia empírica de los autores en la conducción de síntesis de evidencias. **Resultados:** Se propone un flujo metodológico estructurado en cinco etapas: (1) exportación y deduplicación inicial en EndNote; (2) deduplicación adicional y fragmentación del corpus en bloques anuales en Rayyan; (3) creación de un proyecto en Claude AI, incluyendo los criterios de elegibilidad; (4) tamizaje de títulos y resúmenes asistido por inteligencia artificial (IA), con generación automatizada de una hoja de cálculo que contiene la decisión y la justificación para cada registro; y (5) retorno a Rayyan para la revisión humana independiente y su posterior integración al diagrama de flujo PRISMA. Este proceso configura un modelo teórico-operacional que proporciona recomendaciones prácticas para la definición de los criterios de elegibilidad, la elaboración estructurada de prompts y el registro transparente de las decisiones. **Conclusión:** La guía establece una estructura innovadora para la incorporación de la IA en la etapa de selección de estudios en revisiones de literatura. Este enfoque acelera las fases iniciales de la síntesis de evidencias y orienta el uso responsable de la tecnología. Como limitación, se señala la ausencia de validación empírica multicéntrica del flujo propuesto, lo que demanda su aplicación y evaluación en futuras investigaciones.

Descriptores: Inteligencia Artificial; Inteligencia Artificial Generativa; Práctica Clínica Basada en la Evidencia; Metodología como un Tema; Investigación; Revisión Sistemática.

Editores:
Rosimere Ferreira Santana (ORCID: 0000-0002-4593-3715)
Geilsa Soraia Cavalcanti Valente (ORCID: 0000-0003-4488-4912)
Emerson Willian Santos de Almeida (ORCID: 0000-0002-6846-021X)

Editora:
Escola de Enfermagem Aurora de Afonso Costa – UFF
Rua Dr. Celestino, 74 – Centro, CEP: 24020-091 – Niterói, RJ, Brasil
Correo electrónico de la revista: objn.cme@id.uff.br

Autora correspondiente:
Yasmim Ribeiro Fracaroli
Correo electrónico: yasmim.fracaroli@sou.unifal-mg.edu.br

Lo que ya se sabe:

- Los métodos tradicionales de tamizaje en revisiones exigen una alta inversión de tiempo y presentan limitaciones operacionales en el proceso de síntesis de evidencias.
- La IA generativa posee el potencial de acelerar la selección de estudios, por título y resumen, en investigaciones en el área de la salud.
- Todavía existen vacíos metodológicos referentes a la estandarización, la transparencia y la validación del uso de la IA en revisiones sistemáticas.

Lo que este artículo aporta:

- El uso de Claude AI puede reducir el tiempo invertido en el tamizaje inicial de estudios para revisiones de literatura en salud.
- El flujo metodológico propuesto amplía la transparencia y la estandarización de las decisiones durante la selección de estudios asistida por IA.
- La integración entre Claude AI, Rayyan, EndNote y PRISMA favorece una mayor trazabilidad y organización en el proceso de síntesis de evidencias.

INTRODUCCIÓN

Los estudios de revisión son instrumentos centrales en la Práctica Basada en la Evidencia (PBE), ya que permiten reunir y analizar críticamente el conocimiento disponible sobre una temática determinada. De este modo, ofrecen un soporte calificado para la toma de decisiones clínicas y para la formulación de políticas de salud. Tales estudios sintetizan el saber de forma confiable, favorecen la detección de vacíos en el campo investigado y orientan la definición de nuevas preguntas de investigación[1-3].

El tamizaje de artículos mediante la lectura de títulos y resúmenes constituye una de las etapas más exhaustivas y demoradas de la revisión de literatura, con un impacto directo en la validez y la reproducibilidad de los resultados. Las estimaciones apuntan que una revisión sistemática convencional consume, en promedio, de seis a 18 meses de trabajo. La fase de selección de estudios representa una proporción sustancial de este período, volviéndose particularmente vulnerable a la fatiga humana y a la deriva interpretativa entre los revisores[4,5].

Los gestores de referencias, como EndNote, y las plataformas dedicadas al tamizaje colaborativo, como Rayyan, se han consolidado como recursos indispensables en el flujo de las revisiones. Aunque estas tecnologías automatizan la deduplicación y la organización del *corpus* de manuscritos identificados en las búsquedas, la decisión de elegibilidad permanece centrada en la lectura, la interpretación y la cognición humana, mediante la evaluación de cada registro de forma individual[6,7].

La Inteligencia Artificial (IA) ha transformado los estudios de revisión al automatizar etapas como la búsqueda, el tamizaje y la síntesis de datos, confiriendo mayor celeridad y precisión al proceso metodológico[8]. Más recientemente, se destacan los *Large Language Models* (LLMs), sistemas capaces de comprender, generar y estructurar textos complejos, lo que amplía significativamente las posibilidades de soporte tecnológico para la selección de estudios en estas investigaciones[9].

Modelos como ChatGPT (OpenAI), Gemini (Google) y Claude AI (Anthropic) han demostrado, en evaluaciones preliminares, un desempeño prometedor en el tamizaje de títulos y resúmenes. En dominios específicos, estas herramientas alcanzaron una sensibilidad y especificidad comparables a las de los revisores humanos[10,11]. No obstante, la validación humana sigue siendo imperativa para asegurar la calidad metodológica, mitigar sesgos y garantizar interpretaciones contextualizadas[8,12].

Entre las tecnologías disponibles, Claude AI presenta particularidades que lo tornan singularmente atractivo para las revisiones en el área de la salud. La herramienta posee una amplia ventana de contexto (hasta 200 mil *tokens*), soporta la creación de proyectos con archivos persistentes adjuntos y procesa extensos lotes de registros bibliográficos en una sola interacción. Además, genera justificaciones auditables, en lenguaje natural, para cada decisión de inclusión o exclusión de artículos[13]. Tales recursos viabilizan el mantenimiento de criterios de elegibilidad estables a lo largo del tamizaje y confieren trazabilidad al proceso. Estos atributos son fundamentales para la transparencia preconizada por las directrices *Preferred Reporting Items for Systematic reviews and Meta-Analyses* (PRISMA) y por los manuales Cochrane Handbook y JBI[14-16].

A pesar de los avances recientes, persiste un vacío en la literatura en cuanto a la ausencia de orientaciones metodológicas sistematizadas para integrar —de forma estructurada, crítica y ética— el uso de LLMs a la etapa de selección de estudios. La producción científica actual tiende a concentrarse en la comparación de desempeño entre la IA y los revisores humanos o en discusiones teóricas sobre los riesgos de la automatización. En consecuencia, se carece de una guía práctica que articule las etapas intermedias —como la gestión de referencias, la deduplicación, la fragmentación del *corpus* y la configuración de la herramienta— con la decisión final de elegibilidad de los estudios[16-18].

Ante este escenario, este estudio tiene como objetivo presentar una guía metodológica para la selección de estudios en revisiones de literatura con el auxilio de IA generativa —específicamente del modelo Claude AI—, con un enfoque en la optimización de la síntesis de evidencias en el área de la salud.

MÉTODO

Se trata de la proposición de una guía metodológica fundamentada en una reflexión teórico-metodológica. El proceso se condujo mediante una revisión narrativa de la literatura sobre el uso de la IA generativa en la síntesis de evidencias, aliada a la experiencia de los autores en la conducción de diferentes tipos de revisiones (sistemáticas, integrativas y de alcance). El desarrollo de la guía se orientó por las directrices internacionales PRISMA[14], Cochrane Handbook[15] y JBI Manual for Evidence Synthesis[16], además de adoptar los preceptos reconocidos para la elaboración de manuales orientadores en salud[19].

La reflexión ocurrió de forma estructurada en cinco

etapas: (1) levantamiento y análisis crítico de publicaciones metodológicas y de estudios empíricos sobre la aplicación de LLMs en el tamizaje de estudios[8,12,17,18]; (2) mapeo de las funcionalidades de estas herramientas en relación con los requisitos de la directriz PRISMA para el tamizaje de títulos y resúmenes; (3) análisis comparativo entre las recomendaciones metodológicas de organizaciones de referencia internacional (como Cochrane y JBI) y las potencialidades y limitaciones operacionales de Claude AI; (4) sistematización de la experiencia práctica de los autores en la aplicación iterativa del flujo propuesto, con el objetivo de identificar oportunidades, riesgos metodológicos y requisitos para el uso responsable de la herramienta; y (5) consolidación de las etapas identificadas en un flujo metodológico, culminando en la proposición de la guía y del diagrama de flujo operacional expuestos en la sección de Resultados.

La búsqueda en la literatura se realizó en las fuentes de información Medical Literature Analysis and Retrieval System Online (PubMed/MEDLINE), Scopus, Web of Science Core Collection (WoS) y Google Scholar. Se estableció el recorte temporal de 2020 a 2025, justificado por la recencia del lanzamiento comercial de los LLMs. Se utilizaron los descriptores "artificial intelligence", "large language model", "systematic review", "screening", "title and abstract" y "Claude AI", así como sus correspondientes en portugués, cruzados mediante los operadores booleanos AND y OR. Se incluyeron estudios publicados en inglés, portugués o español, sin restricción de diseño metodológico, que abordasen el uso de la IA en la síntesis de evidencias. Se excluyeron los artículos sin una descripción metodológica clara o que divergían del escrutinio temático.

Por tratarse de una revisión narrativa, de carácter reflexivo y no sistemático, se prescindió del registro previo de protocolo en plataformas como el International Prospective Register of Systematic Reviews (PROSPERO) o el Open Science Framework (OSF), de conformidad con la naturaleza del diseño del estudio[17].

El tamizaje de la literatura ocurrió en dos fases: lectura de títulos y resúmenes, seguida de la evaluación en su integridad de los textos potencialmente relevantes. El análisis de los datos se condujo de forma interpretativa, con un foco en la identificación de estrategias, limitaciones y recomendaciones para el uso de la IA en la selección de artículos. Se priorizaron las experiencias aplicables al contexto de la salud, con énfasis en propuestas metodológicas y estudios de validación empírica.

La experiencia empírica de los autores se incorporó de forma sistematizada a partir de la aplicación práctica y de la reflexión crítica sobre el uso de Claude AI durante la conducción de una revisión sistemática enfocada en protocolos de desensibilización a taxanos en mujeres con cáncer de mama. Tal investigación se encuentra en curso y está debidamente registrada en el PROSPERO (CRD420261325732). La elección de este estudio como escenario de aplicación sobrevino de su sinergia con la línea de investigación del grupo, viabilizando la prueba y el refinamiento de la guía en un contexto real de producción científica. Durante la fase de prueba, se procesaron bloques anuales de registros bibliográficos —previamente deduplicados en EndNote y Rayyan—, insertándose los criterios de elegibilidad como archivos de contexto persistentes en un proyecto dedicado en Claude AI. El equipo de investigación analizó críticamente las decisiones y las justificaciones generadas por el modelo, lo que propició el perfeccionamiento iterativo del método aquí documentado.

En lo tocante a la seguridad y a la privacidad de los datos, la aplicación del flujo ocurrió mediante el uso de la versión Claude Pro (interfaz claude.ai). Se desactivó la configuración de privacidad "Help Improve Claude" para impedir que los comandos (*prompts*) y los archivos adjuntos fuesen utilizados en el entrenamiento de modelos futuros de la empresa Anthropic[20]. El suministro de información a la herramienta se restringió a los metadatos bibliográficos de los registros (identificador, título, autores, año, periódico, DOI, resumen y palabras clave). No se sometieron textos en su integridad, datos inéditos o cualesquiera otros materiales sensibles.

La proposición metodológica resultante se estructuró en cinco etapas secuenciales, descritas en la sección de Resultados e ilustradas en un diagrama de flujo operacional. Por tratarse de un estudio de naturaleza teórico-metodológica, sin la involucración directa de seres humanos, se dispensó la apreciación por un Comité de Ética en Investigación, respetándose íntegramente las exigencias de la Ley nº 14.874/2024 y de las Resoluciones nº 466/2012 y nº 510/2016 del Consejo Nacional de Salud (CNS).

RESULTADOS

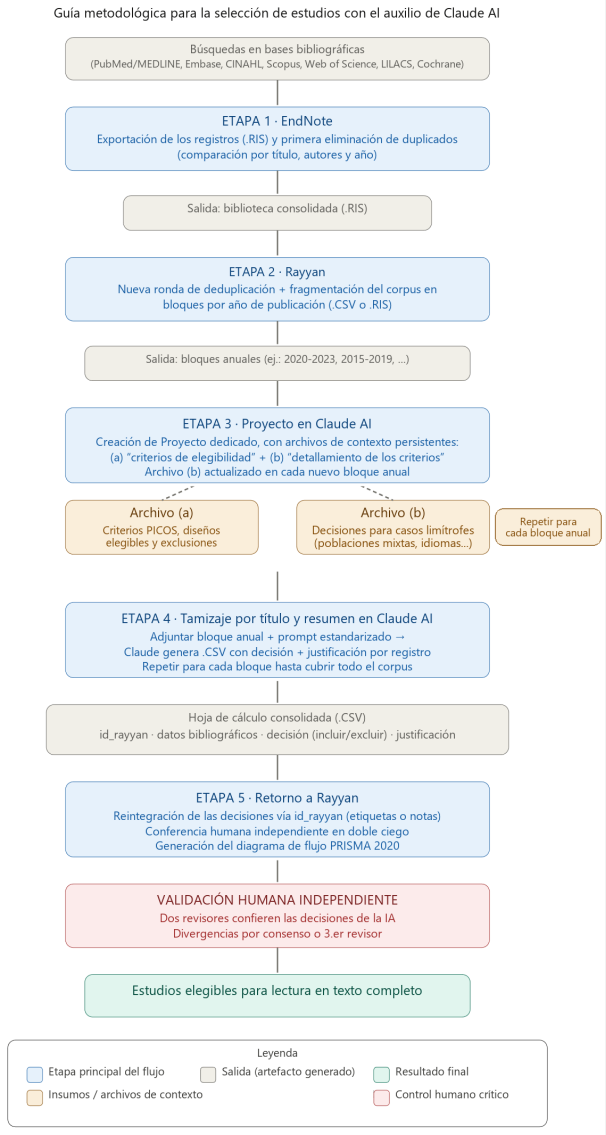
La guía metodológica propuesta se estructura en cinco etapas secuenciales, articulando herramientas tradicionales de gestión de referencias y de tamizaje colaborativo (EndNote y Rayyan) con el uso de Claude AI como recurso auxiliar en la decisión de elegibilidad por título y resumen. La Figura 1 presenta la representación esquemática del flujo.

Etapa 1 – Exportación de las fuentes de información y primera deduplicación en EndNote

Tras la ejecución de las estrategias de búsqueda en las fuentes de información seleccionadas (PubMed/MEDLINE, Embase, LILACS, Scopus, Web of Science y Cochrane Library), todos los registros recuperados deben ser exportados en un formato compatible con el gestor de referencias EndNote. Se recomienda priorizar los archivos con la extensión .RIS o el formato nativo de cada base de datos[5].

En EndNote, los registros procedentes de las diferentes fuentes se consolidan en una única biblioteca, en la cual se realiza la primera ronda de eliminación de duplicados mediante la función "Find Duplicates". Se sugiere configurar el programa para comparar el título, los autores y el año de publicación. Además, es imperativa la revisión manual del conjunto antes de la exclusión definitiva, ya que las variaciones tipográficas y de formato —como la presencia o ausencia del Digital Object Identifier (DOI) y las distintas formas de citación de los autores— pueden generar falsos negativos en el algoritmo automático[5,15]. La salida de esta etapa consiste en una biblioteca consolidada y parcialmente depurada, que deberá ser exportada nuevamente en formato .RIS para la fase subsiguiente.

Figura 1 – Diagrama de flujo operacional de la guía metodológica para la selección de estudios en revisiones de la literatura con el auxilio de Claude AI. Alfenas, MG, Brasil, 2026



Fuente: elaborado por los autores, 2026.

Etapa 2 – Nueva deduplicación y fragmentación en Rayyan

A continuación, la biblioteca es importada a Rayyan, una plataforma inteligente ampliamente utilizada en el tamizaje de artículos para revisiones. La herramienta ofrece una versión gratuita, con funcionalidades restringidas, y planes de pago, con recursos avanzados[6]. Rayyan dispone de un algoritmo propio para la detección y exclusión de duplicados, el cual presenta una sensibilidad superior a la de EndNote en diversos escenarios, lo que justifica la realización de una segunda ronda de deduplicación en esta etapa[6,15].

Concluida la depuración, el corpus consolidado se fracciona en bloques más pequeños, organizados por el año de publicación. La fragmentación anual ofrece ventajas operacionales relevantes: (i) facilita el rastreo temporal de las publicaciones y la auditoría del proceso de tamizaje; (ii) genera archivos de un tamaño compatible con la ventana de contexto de Claude AI, previniendo el truncamiento de datos; y (iii) posibilita el trabajo simultáneo entre los revisores o la reanudación parcial en casos de interrupción. Cada bloque se exporta en los formatos .CSV o .RIS y debe contener, como mínimo, los siguientes campos: identificador único, título, autores, año, periódico, DOI, resumen y palabras clave.

Etapa 3 – Creación del Proyecto en Claude AI y configuración de los criterios de elegibilidad

El recurso de "Proyectos" de Claude AI viabiliza la creación de un espacio de trabajo dedicado, en el cual los archivos de contexto pueden ser adjuntados de forma persistente y permanecen accesibles en todas las interacciones vinculadas a ese entorno[11]. Esta funcionalidad es el núcleo de la estrategia propuesta aquí, pues transforma los criterios de elegibilidad en un referencial decisorio fijo, garantizando la consistencia durante la evaluación de los diferentes bloques documentales.

Se recomienda crear un Proyecto específico para cada revisión e instruirlo con dos documentos complementarios, adjuntados como contexto: a) Documento con los criterios de elegibilidad (versión narrativa): conteniendo los elementos habitualmente descritos en el protocolo de la revisión, como población, intervención, comparación, desenlaces, diseños de estudio elegibles y criterios de exclusión. b) Documento de detalle operacional: explicitando las decisiones interpretativas para el manejo de casos limítrofes. Puede incluir, por ejemplo, la conducta ante estudios con poblaciones mixtas, la definición precisa de la intervención, el tratamiento de las revisiones sistemáticas como fuentes secundarias, la política referente a los idiomas de publicación y demás convenciones acordadas por el equipo.

La segregación entre el documento de criterios y el de detalle operacional —estrategia adoptada con éxito en la experiencia de los autores— presenta dos ventajas centrales. Primero, preserva la integridad del protocolo original (generalmente registrado en plataformas como PROSPERO u OSF). Segundo, permite el perfeccionamiento y el versionamiento de las reglas operacionales a lo largo del tamizaje, sin alterar el documento formalmente registrado. Esto confiere flexibilidad al proceso y resguarda la trazabilidad. La Tabla 1 sintetiza los elementos mínimos recomendados para la composición de estos dos documentos.

Paso a paso para la creación del Proyecto en Claude AI

Para subsidiar la operacionalización de esta etapa, se describe la secuencia de procedimientos para la creación y la configuración del Proyecto en Claude AI, ilustrada en la Figura 2.

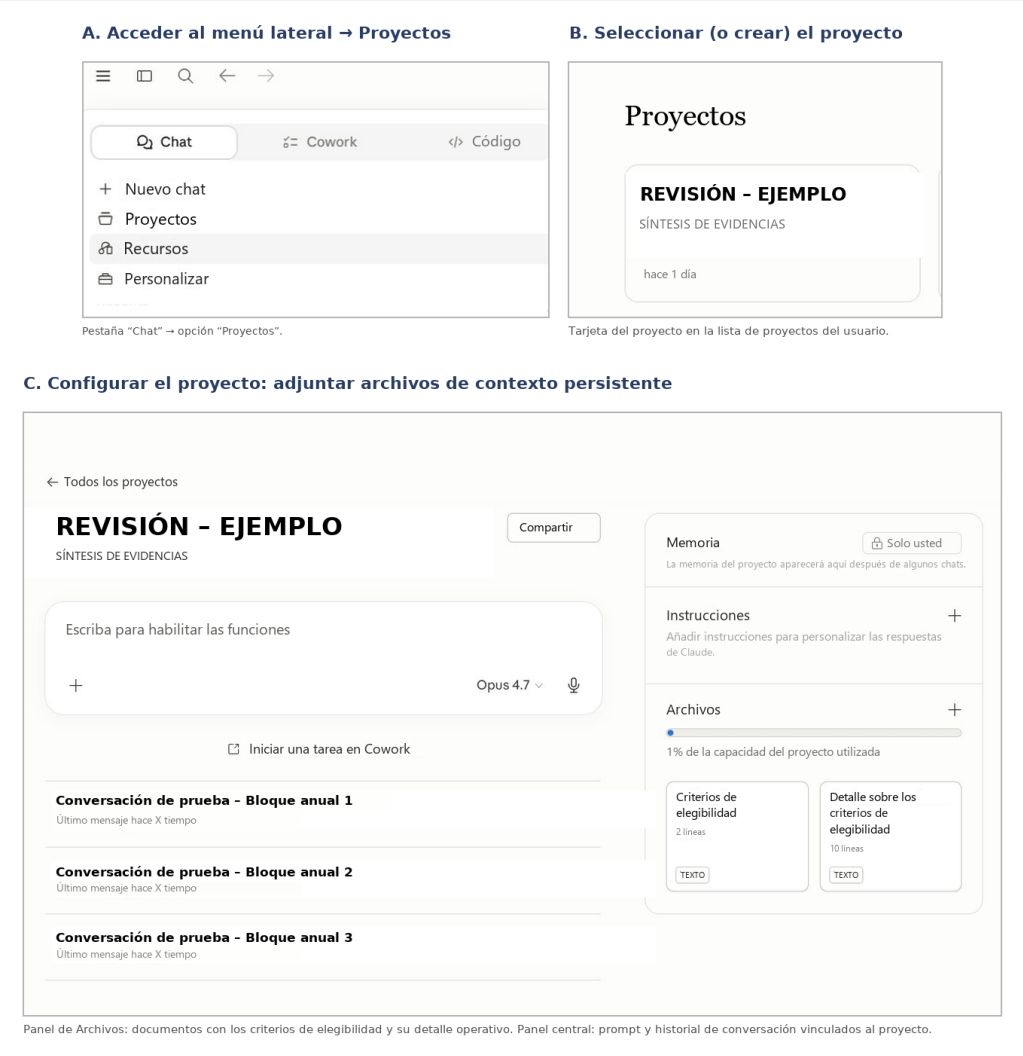
Paso 1 - Acceder al menú lateral y seleccionar "*Projects*": Tras efectuar el registro o el inicio de sesión en la interfaz (https://claude.ai/new), el usuario debe abrir el menú de navegación en el lateral izquierdo y seleccionar la opción "*Projects*" (Figura 2 - Panel A).

Tabla 1 – Elementos mínimos recomendados para la configuración de los criterios de elegibilidad en el Proyecto de Claude AI. Alfenas, MG, Brasil, 2026

Elemento	Contenido esperado
Población	Definición clara de los participantes, con especificación de la edad, el sexo, la condición clínica y los criterios diagnósticos exigidos.
Intervención	Descripción precisa de la intervención, incluyendo la definición operacional (p. ej., número mínimo de etapas, dosis, vías de administración).
Comparación	Grupo de comparación o estándar de referencia adoptado (cuando corresponda).
Desenlaces	Lista de los desenlaces primarios y secundarios considerados elegibles.
Diseños elegibles	Tipos de estudio aceptados (ECA, ENoC, cohorte, caso-control, etc.) y diseños explícitamente excluidos.
Criterios de exclusión	Situaciones que motivan la exclusión inmediata, como un rango de edad fuera del alcance, la presencia de otras patologías y formatos editoriales.
Idiomas y período	Política sobre los idiomas aceptados y la ventana temporal de la búsqueda.
Casos limítrofes	Reglas de decisión para poblaciones mixtas, datos extraíbles por separado, revisiones previas, estudios sin resumen, entre otros escenarios ambiguos.

Fuente: elaborado por los autores, 2026.

Figura 2 – Paso a paso para la creación y la configuración de un Proyecto en Claude AI, orientado al tamizaje asistido en revisiones de la literatura. Alfenas, MG, Brasil, 2026



Fuente: elaborado por los autores, 2026.

Paso 2 - Crear y nombrar un nuevo Proyecto: En la pantalla de listados, se selecciona la opción para crear un nuevo proyecto. Se recomienda utilizar una nomenclatura fácilmente identificable (p. ej., la sigla o el título resumido de la revisión) y añadir una breve descripción que especifique el tipo de síntesis (revisión sistemática, integrativa, de alcance, etc.). Esta estandarización optimiza la localización y la organización del trabajo en equipo (Figura 2 - Panel

B).

Paso 3 - Adjuntar los archivos de contexto persistente: En el panel derecho de la interfaz, en la sección "Archivos", se deben adjuntar los dos documentos balizadores del tamizaje: los criterios de elegibilidad y el detalle operacional. Estos archivos permanecerán disponibles en todas las interacciones, asegurando la aplicación consistente de los mismos parámetros en cada bloque anual evaluado (Figura 2 - Panel C). Adicionalmente, la interfaz posibilita la inserción de directrices generales en el campo "Instrucciones", las cuales orientarán el comportamiento del modelo. El panel "Memoria" registrará automáticamente información relevante acumulada durante las interacciones, fortaleciendo la trazabilidad del método.

Etapa 4 – Tamizaje asistido por título y resumen en Claude AI

Con el Proyecto configurado, se inicia la ejecución del tamizaje. Cada bloque anual de registros se adjunta, individualmente, a una nueva conversación dentro del Proyecto. El comando (*prompt*) introducido debe solicitar, de forma explícita: (i) la lectura de los títulos y resúmenes de todos los ítems del archivo; (ii) la aplicación de los criterios de elegibilidad y del detalle operacional previamente adjuntados; y (iii) la generación de una hoja de cálculo (.CSV) que contenga el identificador del registro, los metadatos bibliográficos esenciales, el veredicto (incluir para lectura en su integridad o excluir) y la respectiva justificación textual.

La justificación individualizada por registro es el elemento central de este método, pues actúa como un mecanismo de auditoría y rendición de cuentas (*accountability*). Al contrario de los sistemas que proporcionan únicamente clasificaciones binarias, Claude AI es capaz de articular, en lenguaje natural, qué criterio exacto balizó la decisión (p. ej., "revisión narrativa sobre taxanos sin la descripción de un protocolo formal de desensibilización" o "estudio preclínico realizado en cultivo celular"). Esta transparencia posibilita la posterior comprobación humana y la mitigación de discordancias.

Se recomienda la descarga y el archivo de las hojas de cálculo generadas para cada bloque. Estas actuarán como artefactos del proceso de selección e integrarán el acervo auditable de la investigación. La Tabla 2 presenta un modelo de *prompt*, basado en la experiencia de los autores y perfeccionado de manera iterativa a lo largo de las pruebas.

Tabla 2 – Modelo de prompt utilizado para el tamizaje de estudios en Claude AI. Alfenas, MG, Brasil, 2026.

Solicitud principal: Realice el tamizaje de los estudios disponibles en el archivo adjunto llamado '[nombre_del_bloque].csv' basado en la lectura de los títulos y resúmenes, conforme a los criterios de elegibilidad establecidos y disponibles en los archivos del proyecto llamados 'criterios de elegibilidad' y 'detalle sobre los criterios de elegibilidad'". **Salida esperada:** Produzca un nuevo archivo .csv que contenga, para cada registro analizado, las siguientes columnas: (1) id_triagem — identificador secuencial atribuido en el tamizaje; (2) id_rayyan (o key) — identificador original del registro en Rayyan; (3) título del estudio; (4) año de publicación; (5) periódico; (6) volumen, número y páginas; (7) autores; (8) DOI; (9) idioma; (10) resumen; (11) palabras clave; (12) PubMed ID (cuando esté disponible); (13) triage_decision — 'Incluir para ver el texto completo' o 'Excluir'; (14) justificativa — explicación textual fundamentada en el criterio específico de elegibilidad o exclusión que motivó la decisión. **Instrucciones adicionales:** • Adopte una postura conservadora: en caso de duda sobre el cumplimiento de los criterios, marque el registro como 'Incluir para texto completo' y justifique la incertidumbre. • Para los registros sin resumen, base la decisión en el título y en las palabras clave, indicando esta limitación en la justificación. • Mantenga la codificación 8-bit Unicode Transformation Format (UTF-8) (con BOM) y el separador de coma en el archivo .csv generado, con todos los campos entre comillas para preservar el contenido que contenga comas y puntos y comas".

Fuente: elaborado por los autores, 2026.

Etapa 5 – Retorno al Rayyan para la conferencia e integración al PRISMA

Tras la finalización del tamizaje por Claude AI, es imprescindible reintegrar las decisiones al entorno de Rayyan. Este procedimiento permite: (i) la ejecución de la conferencia humana independiente; (ii) la consolidación de todos los datos en un entorno auditable único; y (iii) la generación automatizada del diagrama de flujo PRISMA, contemplando el recuento preciso de las inclusiones y exclusiones en cada fase metodológica.

Operacionalmente, las hojas de cálculo anuales generadas por la IA deben compilarse en una única hoja de cálculo maestra. La columna id_rayyan debe mantenerse como clave primaria, pues posibilita la localización exacta de cada registro importado originalmente. Se recomienda la comprobación rigurosa de estos identificadores antes del reingreso a la plataforma, con el fin de subsanar eventuales

duplicidades o fallas de codificación. En Rayyan, las decisiones de la IA pueden aplicarse mediante la atribución de etiquetas (*labels*) personalizadas (p. ej., 'IA-incluir' e 'IA-excluir'). Este enfoque preserva el veredicto de la herramienta en una capa informacional distinta de la generada por los investigadores. Otra estrategia plausible es la inserción de las justificaciones generadas por el modelo en el campo de notas (*notes*) de cada registro, optimizando el análisis comparativo. La elección del método debe ponderar el volumen del *corpus*, el tamaño del equipo de revisores y las preferencias del grupo de investigación.

La validación humana independiente —conducida por dos revisores en formato doble ciego en Rayyan— debe evaluar la integridad de las decisiones emitidas por la IA. Las discordancias entre el modelo y el investigador, o entre los propios investigadores, deben dirimirse por consenso o mediante el dictamen de un tercer revisor, alineándose con los preceptos del *Cochrane Handbook* y de JBI[15,16]. Este rigor metodológico consagra a la IA en el papel de primer revisor automatizado, manteniendo la ineludible supervisión humana, garante de la calidad de las evidencias sintetizadas.

Finalizada la validación, Rayyan posibilita la exportación cuantitativa de los estudios identificados, deduplicados, incluidos y excluidos, proporcionando los insumos para la confección del diagrama de flujo PRISMA.

Se recomienda, por último, que el informe final de la revisión documente expresamente el uso de la IA generativa en la etapa de tamizaje de títulos y resúmenes. Se debe incluir una nota metodológica que detalle los procedimientos de validación adoptados, en estricta obediencia a las directrices para el reporte del empleo de inteligencia artificial en estudios de revisión[21].

Transparencia y registro de la experiencia de aplicación

La revisión sistemática utilizada como escenario práctico para la prueba del flujo propuesto se encuentra registrada en la base PROSPERO (CRD420261325732) y está disponible para consulta pública. Los resultados cuantitativos completos de este tamizaje —incluyendo el número de registros identificados, deduplicados, tamizados por Claude AI y confirmados en la conferencia humana independiente, además del respectivo diagrama de flujo PRISMA— se detallarán en la publicación primaria del estudio. Esta revisión se encuentra actualmente en desarrollo por el mismo grupo de investigación. La opción de no presentar los datos cuantitativos en este momento busca evitar la publicación duplicada de datos y preservar la originalidad de la síntesis de evidencias para su artículo principal.

Se destaca que el objetivo central de esta guía es describir y fundamentar el flujo operacional elaborado, y no reportar los hallazgos de una revisión sistemática específica. Tales resultados serán oportunamente publicitados en un manuscrito propio, con la debida citación cruzada a este estudio metodológico.

Ante la creciente adopción de herramientas de IA en el apoyo a la selección y al tamizaje de estudios, la Tabla 3 presenta un análisis comparativo entre el tamizaje manual, el uso de Claude AI y la aplicación de otras tecnologías. La tabla evidencia las principales diferencias relacionadas con los aspectos operacionales, contextuales y funcionales.

Tabla 3 – Comparación entre el tamizaje manual, Claude AI y otras herramientas de IA en el apoyo a la selección de estudios. Alfenas, MG, Brasil, 2026

Aspecto	Tamizaje manual	Claude AI	Otras IAs (p. ej., ChatGPT, Gemini, ASReview)
Tiempo medio por bloque (~400 registros)	De 8 a 15 horas por revisor.	De minutos a pocas horas, según el tamaño del archivo.	Variable. Las herramientas con aprendizaje activo (como ASReview) reducen el esfuerzo de forma progresiva.
Consistencia entre lotes	Susceptible a la fatiga y a la deriva interpretativa a lo largo del tiempo.	Alta, siempre que los archivos de contexto se mantengan inalterados en el Proyecto.	Depende de la memoria de la sesión. ChatGPT y Gemini exigen una reconfiguración en cada nueva conversación.
Justificación por registro	Cuando se registra, frecuentemente es sucinta.	Generada de manera automática en lenguaje natural y vinculada a los criterios de elegibilidad.	ASReview entrega una puntuación de probabilidad; ChatGPT y Gemini generan texto libre.
Ventana de contexto	No se aplica.	Extensa; soporta amplios lotes de registros sin truncamiento.	ChatGPT y Gemini poseen ventanas competitivas; ASReview no utiliza LLMs.
Criterios persistentes	El revisor consulta el documento del protocolo.	Los archivos de contexto permanecen fijados al entorno del Proyecto.	Se exige pegar o volver a adjuntar los criterios de forma manual en cada nueva interacción.
Costo	Tiempo de trabajo del investigador.	Suscripción mensual; uso directo a través de la interfaz claude.ai.	Variable. ASReview es gratuito; otras soluciones comerciales cobran por uso o suscripción.
Auditabilidad	Alta, en caso de que exista un registro detallado de las decisiones por parte de los revisores.	Alta (genera una hoja de cálculo estandarizada que contiene el veredicto y la justificación para cada registro).	Variable, dependiendo de la arquitectura y de las funcionalidades de la herramienta utilizada.

Fuente: elaborado por los autores, 2026.

DISCUSIÓN

El presente estudio no tuvo como propósito el desarrollo de una nueva tecnología de IA, tampoco la configuración de un agente autónomo capaz de operar sin la supervisión humana. Claude AI se utilizó en su forma nativa, mediante *prompts* dirigidos manualmente por los investigadores dentro de un Proyecto dedicado, actuando como un modelo de lenguaje asistido, y no como un agente de ejecución independiente. La innovación propuesta en este estudio reside, por lo tanto, en la articulación sistematizada de un flujo operacional replicable. Este flujo integra herramientas ya consolidadas (EndNote y Rayyan) con el uso estructurado de LLMs comerciales, fundamentándose en un conjunto original de recomendaciones prácticas. Tales recomendaciones incluyen: la separación entre los criterios formales de elegibilidad y el detalle operacional para casos limítrofes; la estandarización de *prompts* con una estructura de salida auditable; el uso de identificadores únicos (id_rayyan) como mecanismo de trazabilidad entre los entornos virtuales; y la aplicación de un protocolo de conferencia humana independiente tras el tamizaje por la IA.

A diferencia de estudios anteriores —centrados en la comparación de desempeño entre la IA y los revisores humanos o en discusiones teóricas sobre los riesgos de la automatización[8,12,18]—, la presente guía ofrece un modelo teórico-operacional probado empíricamente en una revisión sistemática en curso. De esta forma, transforma el uso *ad hoc* de los modelos de lenguaje en un proceso transparente, auditable y alineado con las directrices internacionales de síntesis de evidencias. Además, el flujo atiende a las recomendaciones específicas para el reporte transparente del uso de la IA como herramienta metodológica en revisiones de literatura[21].

En este sentido, el método propuesto debe analizarse bajo tres prismas interdependientes: como estrategia de eficiencia operacional, como intervención metodológica con riesgos inherentes y, fundamentalmente, como vector de impacto transformador en la práctica científica en salud. La incorporación de la IA en la etapa de selección de estudios reconfigura el trabajo metodológico de las revisiones, desplazando el esfuerzo operacional hacia los sistemas computacionales. Esto libera una mayor inversión cognitiva de los revisores para la inferencia, la interpretación y la evaluación crítica de las decisiones analíticas. Esta transformación, sin embargo, introduce nuevas dependencias técnicas —como el modelo elegido, su versión y sus parámetros—, además de nuevas fuentes de variación que exigen una gobernanza rigurosa para mantener la trazabilidad y la responsabilidad científica[9].

Potencialidades de Claude AI en la optimización de la selección de estudios

En el plano de la eficiencia, Claude AI demuestra un alto potencial al desplazar el esfuerzo operacional del tamizaje, por título y resumen, hacia los sistemas computacionales. Estudios recientes cuantifican esta ventaja al evidenciar que las herramientas basadas en LLMs pueden alcanzar una precisión y recuperación comparables a las de los revisores humanos, con una reducción significativa del tiempo necesario para la selección de los manuscritos. Los análisis comparativos demuestran puntuaciones F1 (F1-scores) superiores al 90 % en las tareas de tamizaje, con la reducción del tiempo de evaluación de días a horas en el análisis de centenares de registros bibliográficos[10,22].

Esta optimización permite la reasignación de los recursos humanos hacia las etapas de mayor exigencia cognitiva de la investigación, como el análisis crítico de los hallazgos y la síntesis cualitativa. Adicionalmente, la extensa ventana de contexto de Claude AI y el recurso de Proyectos con archivos persistentes viabilizan el procesamiento de amplios lotes de registros sin la fragmentación artificial del trabajo, configurando un diferencial operacional relevante en relación con otras herramientas de IA[13].

Considerando la creciente adopción de estas tecnologías en el apoyo a la selección y al tamizaje de artículos, la Tabla 3 (presentada en la sección de Resultados) propone un análisis comparativo entre Claude AI y otros sistemas aplicables a la misma finalidad, evidenciando las diferencias relacionadas con los aspectos operacionales, contextuales y funcionales. Los datos sintetizados demuestran que, aunque plataformas como ASReview presenten ventajas específicas en costo y en aprendizaje activo progresivo, Claude AI se destaca por la combinación de tres factores: una amplia ventana de contexto, la persistencia de los criterios a lo largo de todo el tamizaje y la generación de justificaciones en lenguaje natural vinculadas a cada decisión. En conjunto, estos atributos favorecen sustancialmente la auditabilidad del proceso metodológico.

Desafíos éticos y metodológicos

El impacto metodológico de la IA sobre la validez y la confiabilidad de la selección de estudios exige cautela. Aunque la automatización reduce los errores humanos derivados de la fatiga o la inconsistencia, introduce riesgos sistémicos que demandan una gestión rigurosa. Se destacan los sesgos algorítmicos —notablemente los de selección, medición y confusión—, los cuales pueden desplazar el centro de gravedad de la síntesis hacia escenarios más compatibles con el corpus de entrenamiento del modelo. Este desvío es capaz de generar asimetrías y clasificaciones conceptualmente frágiles[23-25].

En el contexto de la selección de artículos, el sesgo de selección puede manifestarse cuando los modelos entrenados en bases de datos con predominio de determinados idiomas, diseños o regiones geográficas presentan un desempeño inferior al evaluar investigaciones de contextos subrepresentados, comprometiendo la equidad de la síntesis[23]. Sumado a esto, el fenómeno de las "alucinaciones" (generación de información plausible, pero fácticamente incorrecta) representa un riesgo epistemológico singular. En registros sin resumen o con *abstracts* poco informativos, la IA puede producir decisiones aparentemente bien fundamentadas, pero basadas en inferencias no verificables en el texto fuente[22,26].

La ocurrencia de falsos positivos (inclusión indebida de estudios no elegibles) y de falsos negativos (exclusión indebida de estudios elegibles) constituye una preocupación central en cualquier estrategia de tamizaje asistida

por IA, pues impacta directamente en la completitud y la validez de la síntesis final. En la experiencia práctica que fundamentó esta guía, este riesgo se mitigó a través de dos frentes: la orientación explícita para que el modelo adoptara una postura conservadora ante las incertidumbres (priorizando la inclusión para la lectura en su integridad en detrimento de la exclusión precoz) y la conferencia humana independiente de la totalidad de las decisiones generadas. Este enfoque desplaza el riesgo predominante hacia los falsos positivos, considerados metodológicamente menos nocivos en esta etapa, puesto que serán rectificados durante la lectura del texto completo. Los falsos negativos, por otra parte, comprometerían irreversiblemente el alcance de la síntesis.

La confiabilidad de las decisiones, a su vez, depende intrínsecamente de la calidad y de la estabilidad de los prompts utilizados. Los comandos genéricos o inconsistentes entre diferentes ejecuciones pueden generar una variabilidad indebida en el tamizaje, aun cuando los criterios de elegibilidad permanezcan inmutables. La estandarización del prompt y su fijación como elemento auditable del proceso, conforme a lo propuesto en esta guía, se configuran no solo como buenas prácticas operacionales, sino como requisitos metodológicos esenciales para garantizar la reproducibilidad del tamizaje asistido por IA[21].

Para mitigar estos riesgos en el flujo establecido, se recomienda la incorporación de mecanismos de contención. Entre ellos se destacan: la instrucción explícita para que el modelo limite sus inferencias a los datos contenidos en el título y en el resumen; la exigencia de una justificación textual vinculada a un criterio específico de elegibilidad; y la verificación cruzada por revisores humanos independientes. La transparencia y la reproducibilidad son pilares innegociables. Por lo tanto, se exige que la comunidad científica adopte prácticas de ciencia abierta, documentando meticulosamente el modelo de IA, su versión, el prompt y los archivos de contexto utilizados[23,27].

Seguridad, privacidad y transparencia en el uso de la IA

El empleo de herramientas de IA generativa en procesos de investigación impone consideraciones apremiantes sobre la seguridad y la confidencialidad de los datos, sobre todo al manipular metadatos de producción científica inédita o protocolos de revisión en curso. En la experiencia de aplicación de esta guía, la desactivación de la configuración "*Help Improve Claude*" en la versión Claude Pro se configuró como una medida esencial. El objetivo fue impedir que los registros bibliográficos y los criterios de elegibilidad fueran incorporados a los datos de entrenamiento de versiones futuras de la herramienta[20].

Cabe resaltar que dicha desactivación inhibe únicamente el uso de las conversaciones para el entrenamiento del algoritmo, sin eliminar el almacenamiento de los datos en los servidores de la empresa Anthropic durante el período de retención estipulado en su política de privacidad. Asimismo, la versión de consumidor no ofrece las garantías proporcionadas por planes corporativos (como *Claude for Work*), que cuentan con contratos rigurosos de protección, o por implementaciones mediante la interfaz de programación de aplicaciones (*Application Programming Interface* - API) configuradas con retención cero de datos. Esta distinción es crucial en revisiones que abarcan datos sensibles, poblaciones vulnerables o información no publicada. En estos escenarios, la adopción de entornos con salvaguardas contractuales robustas de confidencialidad es sobremanera más adecuada.

En el flujo metodológico propuesto aquí, la restricción de la entrada de datos a metadatos bibliográficos no confidenciales (título, resumen, autores, año, DOI y palabras clave) reduce la exposición de información sigilosa. Con todo, no oblitera integralmente el riesgo, puesto que los resúmenes de revisiones en curso pueden, eventualmente, revelar orientaciones estratégicas de investigaciones aún no divulgadas.

Se recomienda, por lo tanto, que los investigadores evalúen rigurosamente la naturaleza y la sensibilidad de los datos antes de la sumisión a la herramienta. Es imperativo observar la política de privacidad vigente —susceptible a actualizaciones— y constatar la adecuación del plan contratado a la naturaleza del estudio, adoptando el nivel de protección contractual compatible con la información procesada[20].

Limitaciones del estudio

A pesar de su relevancia, el presente estudio posee limitaciones metodológicas. La principal de ellas reside en la ausencia de validación cuantitativa multicéntrica del flujo propuesto. La aplicación práctica se concentró en un único dominio temático, lo que imposibilita estimar con precisión el desempeño del método bajo variaciones contextuales relevantes, tales como diferentes áreas del conocimiento, terminologías singulares, distintos idiomas de publicación y tipologías variadas de síntesis. Asimismo, la eficacia de Claude AI es intrínsecamente dependiente del modelo y de la versión empleados, cuyas actualizaciones futuras pueden modificar los resultados. Tal dinamismo impone la necesidad imperiosa de una documentación técnica granular para asegurar la reproducibilidad metodológica[28].

Implicaciones para la investigación y la práctica en salud

Para la investigación en salud —área marcada por un volumen exponencial de publicaciones y por la premura de respuestas rápidas para orientar la práctica clínica y las políticas públicas—, la adopción de flujos automatizados representa un avance estratégico. El investigador, no obstante, debe asumir el papel de curador crítico de la información. Claude AI actúa como catalizador de la etapa operacional, mientras que la experiencia clínica y metodológica del profesional se resguarda para la interpretación de contextos complejos, la aprehensión de los matices del cuidado y la evaluación de la aplicabilidad de las evidencias en los servicios de salud (con énfasis en el Sistema Único de Salud brasileño, en el contexto local).

En este nuevo escenario tecnológico, la validación humana trasciende el mero control de calidad para consubstanciarse como parte orgánica del proceso metodológico. El papel del revisor se reconfigura: en vez de realizar el ta-

mizaje duplicado exhaustivo de todo el acervo, su atención converge en la reconciliación de las discrepancias entre las decisiones de la IA y la evaluación humana inicial. El esfuerzo cognitivo es, entonces, dirigido hacia el análisis crítico y hacia la interpretación contextualizada de las evidencias[22].

Adicionalmente, el proceso de elaboración y configuración de los criterios de elegibilidad en el Proyecto fomenta una reflexión profunda sobre el protocolo de revisión. Frecuentemente, esta etapa revela ambigüedades operacionales que, de otro modo, emergerían solo en el transcurso del tamizaje. Este atributo confiere al método una dimensión pedagógica sustancial, de particular valor en la formación continuada de investigadores en síntesis de evidencias.

Direcciones futuras

Las futuras investigaciones deben concentrar esfuerzos en el desarrollo de directrices metodológicas específicas para el empleo de la IA en la selección de estudios, estableciendo estándares mínimos de reporte (especificación del modelo, versión, prompt, parámetros y pistas de auditoría). Se impone también la conducción de estudios de validación externa multicéntrica. Tales investigaciones deben confrontar el desempeño de la IA con el de los revisores humanos en dominios temáticos heterogéneos, valiéndose de métricas precisas de sensibilidad y especificidad, además de evaluar el impacto concreto de la herramienta en las conclusiones finales de la síntesis[28,29]. La sedimentación de estas directrices instituirá los cimientos necesarios para el uso responsable, ético y auditable de la IA generativa en la producción de evidencias en salud.

CONCLUSIÓN

La incorporación de Claude AI —un LLM desarrollado por la empresa Anthropic— en la etapa de selección de estudios en revisiones de literatura moderniza el proceso de síntesis de evidencias, ya que automatiza una de las fases más morosas de la producción científica: el tamizaje por título y resumen. Este estudio, de naturaleza teórico-metodológica, presenta una propuesta de flujo estructurado en cinco etapas para orientar el uso responsable de esta tecnología. El enfoque integra el rigor científico a los recursos de Procesamiento de Lenguaje Natural (PLN), demostrando que la automatización puede ser adoptada con seguridad mediante el establecimiento de protocolos claros, una supervisión humana ineludible y una validación continua.

A pesar de la elevada capacidad de Claude AI para interpretar el lenguaje científico y aplicar los criterios de elegibilidad —cuando se configuran como contexto persistente en un Proyecto dedicado—, su utilización demanda cautela en virtud de los riesgos inherentes de los sesgos algorítmicos y las "alucinaciones". Esta salvedad es particularmente crítica en revisiones destinadas a fundamentar decisiones clínicas o a balizar políticas de salud. Por lo tanto, el uso ético y responsable de la herramienta impone la garantía de transparencia, trazabilidad y reproducibilidad en todas las etapas metodológicas.

En términos de implicaciones prácticas, se recomienda que los investigadores: (1) configuren el Proyecto en Claude AI con archivos de contexto persistentes, englobando los criterios de elegibilidad y su respectivo detalle operacional, con el fin de asegurar la consistencia analítica entre los bloques de tamizaje; (2) utilicen comandos (prompts) estructurados y validados de forma iterativa, exigiendo explícitamente una justificación textual para cada registro evaluado; (3) realicen la deduplicación previa en los programas EndNote y Rayyan, fragmentando el corpus bibliográfico en bloques anuales para optimizar el procesamiento y viabilizar la auditoría; y (4) preserven la validación humana independiente (en duplicata) sobre las decisiones generadas por la Inteligencia Artificial, solucionando eventuales divergencias por consenso o por intermedio del dictamen de un tercer revisor.

CONFLICTO DE INTERESES

Los autores declaran no tener conflicto de intereses.

USO DE INTELIGENCIA ARTIFICIAL

Claude AI se utilizó como apoyo metodológico y operativo durante la preparación del artículo, incluyendo la asistencia en la organización del contenido y la generación de las imágenes presentadas en el manuscrito. ChatGPT se utilizó como herramienta auxiliar para la revisión ortográfica, la mejora de la fluidez del texto y el perfeccionamiento de la redacción científica. Además, todas las decisiones metodológicas, los análisis, las interpretaciones, la organización de los datos y el contenido científico del estudio fueron realizados y validados íntegramente por los autores.

REFERENCIAS

1. Farhat N, Tsaioun K, Saunders-Hastings P, Morgan RL, Ramoju S, Hartung T, et al. Systematic review in evidence-based risk assessment. ALTEX. 2022;39(3):463-79. https://doi.org/10.14573/altex.2004111. PMID: 34585732.
2. Wu Q, Nabi G. Systematic reviews and data analyses: guide to clinical practice and research. Scott Med J. 2023;68(3):69-71. https://doi.org/10.1177/00369330231193888. PMID: 37605529.
3. Canto GDL. Revisões sistemáticas da literatura: guia prático. Curitiba: Brazil Publishing; 2020.
4. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 2010;7(9):e1000326. https://doi.org/10.1371/journal.pmed.1000326. PMID: 20877712.
5. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. 2017;7(2):e012545. https://doi.org/10.1136/bmjopen-2016-012545. PMID: 28242767.

6. Bramer WM, Giustini D, Kramer BMR, Anderson PF. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. Syst Rev. 2013;2:115. https://doi.org/10.1186/2046-4053-2-115. PMID: 24360284.
7. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan - a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. https://doi.org/10.1186/s13643-016-0384-4. PMID: 27919275.
8. Ge L, Agrawal R, Singer M, Kannapiran P, Molina JADC, Teow KL, et al. Leveraging artificial intelligence to enhance systematic reviews in health research: advanced tools and challenges. Syst Rev. 2024;13(1):269. https://doi.org/10.1186/s13643-024-02682-2. PMID: 39456077.
9. Maity S, Saikia MJ. Large language models in healthcare and medical applications: a review. Bioengineering (Basel). 2025;12(6):631. https://doi.org/10.3390/bioengineering12060631. PMID: 40564447.
10. Guo E, Gupta M, Deng J, Park YJ, Paget M, Naugler C. Automated paper screening for clinical reviews using large language models: data analysis study. J Med Internet Res. 2024;26:e48996. https://doi.org/10.2196/48996. PMID: 38214966.
11. Syriani E, David I, Kumar G. Screening articles for systematic reviews with ChatGPT. J Comput Lang. 2024;80:101287. https://doi.org/10.1016/j.cola.2024.101287
12. van Dijk SHB, Brusse-Keizer MGJ, Bucsán CC, van der Palen J, Doggen CJM, Lenferink A. Artificial intelligence in systematic reviews: promising when appropriately used. BMJ Open. 2023;13(7):e072254. https://doi.org/10.1136/bmjopen-2023-072254. PMID: 37419641.
13. Claude [Internet]. San Francisco: Anthropic; 2026 [citado 2026 Mayo 18]. Disponible en: https://claude.ai/
14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71. PMID: 33782057.
15. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 6.5 (updated August 2024). London: Cochrane; 2024 [citado 2026 Mayo 18]. Disponible en: https://www.cochrane.org/handbook
16. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis [Internet]. Adelaide: JBI; 2024 [citado 2026 Mayo 18]. Disponible en: https://synthesismanual.jbi.global/. https://doi.org/10.46658/JBIMES-24-01
17. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs. 2005;52(5):546-53. https://doi.org/10.1111/j.1365-2648.2005.03621.x. PMID: 16268861.
18. Wang S, Scells H, Koopman B, Zuccon G. Can ChatGPT write a good Boolean query for systematic review literature search? [Internet]. In: Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23); 2023 Jul 23-27; Taipei, Taiwan. New York (NY): ACM; 2023 [citado 2026 Mayo 18]. p. 1426-36. Disponible en: https://dl.acm.org/doi/10.1145/3539618.3591703. https://doi.org/10.1145/3539618.3591703
19. Echer IC. Elaboração de manuais de orientação para o cuidado em saúde. Rev Lat Am Enfermagem. 2005;13(5):754-7. https://doi.org/10.1590/S0104-11692005000500022. PMID: 16308635.
20. How do I change my model improvement privacy settings? [Internet]. San Francisco: Anthropic; 2026 [citado 2026 Ago 23]. Disponible en: https://privacy.claude.com/en/articles/12109829-how-do-i-change-my-model-improvement-privacy-settings
21. Holst D, Moenck K, Koch J, Schmedemann O, Schüppstuhl T. Transparent reporting of AI in systematic literature reviews: development of the PRISMA-trAIce checklist. JMIR AI. 2025;4:e80247. https://doi.org/10.2196/80247. PMID: 41370833.
22. Khraisha Q, Put S, Kappenberg J, Warraitch A, Hadfield K. Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. Res Synth Methods. 2024;15(4):616-26. https://doi.org/10.1002/jrsm.1715. PMID: 38484744.
23. Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, et al. Fairness of artificial intelligence in healthcare: review and recommendations. Jpn J Radiol. 2024;42(1):3-15. https://doi.org/10.1007/s11604-023-01474-3. PMID: 37540463.
24. Nazer LH, Zatarah R, Waldrip S, Chen Ke JX, Moukheiber M, Khanna AK, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digit Health. 2023;2(6):e0000278. https://doi.org/10.1371/journal.pdig.0000278. PMID: 37347721.
25. Chen F, Wang L, Hong J, Jiang J, Zhou L. Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models. J Am Med Inform Assoc. 2024;31(5):1172-83. https://doi.org/10.1093/jamia/ocae060. PMID: 38520723.
26. Dhuliawala S, Komeili M, Xu J, Raileanu R, Li X, Celikyilmaz A, et al. Chain-of-verification reduces hallucination in large language models [Internet]. In: Findings of the Association for Computational Linguistics: ACL 2024. Stroudsburg: Association for Computational Linguistics; 2024 [citado 2026 Ago 23]. p. 3563-78. Disponible en: https://aclanthology.org/2024.findings-acl.212/. https://doi.org/10.18653/v1/2024.findings-acl.212
27. Ioannidis JP. Transparency, bias, and reproducibility across science: a meta-research view. J Clin Invest. 2024;134(22):e181923. https://doi.org/10.1172/JCI181923. PMID: 39545412.
28. Siontis GCM, Sweda R, Noseworthy PA, Friedman PA, Siontis KC, Patel CJ. Development and validation pathways of artificial intelligence tools evaluated in

randomised clinical trials. BMJ Health Care Inform. 2021;28(1):e100466. https://doi.org/10.1136/bmjhci-2021-100466. PMID: 34969668.

29. Yuan H. Toward real-world deployment of machine learning for health care: external validation, continual monitoring, and randomized clinical trials. Health Care Sci. 2024;3(5):360-64. https://doi.org/10.1002/hcs2.114. PMID: 39479276.

CONTRIBUCIÓN DE AUTORÍA

Concepción del estudio: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Adquisición de datos: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Análisis de datos: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Interpretación de datos: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Todos los autores son responsables de la redacción y revisión crítica del contenido intelectual, de la versión final publicada y de todos los aspectos éticos, legales y científicos relacionados con la exactitud e integridad del estudio.

© 2026 Autores. Los autores conservan los derechos de autor y conceden a la Online Brazilian Journal of Nursing el derecho de primera publicación. Este artículo está licenciado bajo la licencia Creative Commons Atribución 4.0 Internacional (CC BY 4.0).